حریری، نجلا؛ طاهری، مهدی (۱۳۹۳). بررسی کاربرد روش "جزیره داده‌ها" در تولید پیشینه‌های فراداده‌ای با قابلیت نمایه پذیری و پیدانمایی نام‌های برچسب عناصر در محیط موتورهای کاوش وب: بهبود دسترسی به اشیای محتوایی *پژوهش‌نامه کتابداری و اطلاع‌رسانی*، ۴ (۱)، ۹۲- ۷۳.

بررسی کاربرد روش "جزیره داده‌ها" در تولید پیشینه‌های فراداده‌ای با قابلیت نمایه پذیری و پیدانمایی نام‌های برچسب عناصر در محیط موتورهای کاوش وب: بهبود دسترسی به اشیای محتوایی

دکتر نجلا حریری [1]، دکتر سید مهدی طاهری [2]

تاریخ دریافت: ۹۱/۷/۱۷ تاریخ پذیرش: ۹۱/۸/۲۴

چکیده

هدف: *پژوهش حاضر با هدف تبیین کاربرد روش جزیره داده‌ها در تولید پیشینه‌های فراداده‌ای مبتنی بر طرح‌های فراداده‌ای هسته دوبلین، مارک ۲۱، و طرح فراداده‌ای توصیف شیء در قالب زبان نشانه‌گذاری گسترش‌پذیر با قابلیت نمایه‌پذیری و پیدانمایی نام‌های برچسب عناصر در محیط موتورهای کاوش انجام شد.*

روش: *پژوهش با روش تجربی انجام شد.جامعه پژوهش را ۶۰۰پیشینه فراداده‌ای در قالب دو گروه گواه و آزمون تشکیل می‌دادند.گروه گواه دارای ۳۰۰ پیشینه، کاملاً مبتنی بر فرانماهای استاندارد طرح‌های فراداده‌ای مورد مطالعه، و گروه آزمون، مبتنی بر فرانماهای یاد شده اما با ویژگی‌های خاص، دارای ۳۰۰ پیشینه جاسازی شده در برچسب <pre> زبان نشانه‌گذاری فرا متن بر اساس شیوه "جزیره داده‌ها"بودند.*

یافته‌ها: *یافته‌ها حاکی از آن است که نام‌های برچسب تمامی عناصر فراداده‌ای مربوط به پیشینه‌های عضو گروه آزمون که بر اساس روش جزیره داده‌ها تولیدشده بودند، توسط موتورهای کاوش گوگل و یاهو نمایه شده، و در نتایج جست‌وجو پدیدار می‌شوند .اما گروه گواه فاقد این ویژگی بود.بر خلاف گروه گواه، امکان بازیابی پیشینه‌های فراداده‌ای گروه از طریق نام‌های برچسب آن‌ها در موتورهای کاوش وجود دارد.*

کلیدواژه‌ها: *پیدا نمایی، روش جزیره داده‌ها، طرح‌های فراداده‌ای، موتورهای کاوش، نام‌های برچسب عناصر، نمایه‌سازی*

---

۱. دانشیار گروه علم اطلاعات و دانش شناسی دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران ، nadjlahariri@gmail.com

۲. دانش آموخته مقطع دکترای علم اطلاعات و دانش شناسی دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران، taherismster@gmail.com

مقدمه و بیان مسئله

فراداده، ابزار سازماندهی دانش در محیط و رسانه‌های اطلاعاتی جدید، به ویژه وب است. طرح‌های فراداده‌ای که مجموعه‌ای از عناصر با ساختار معناشناختی خاص هستند که به منظور توصیف، شناسایی، کشف، نگهداری، و مدیریت اشیای محتوایی در محیط جدید گسترش یافته‌اند. طرح‌های فراداده‌ای که بر کارکردهای شناسایی و کشف اشیای محتوایی تأکید می‌کنند، فراداده‌های توصیفی نام گرفته‌اند. سازماندهی و توصیف اشیای محتوایی وب بر اساس این نوع فراداده انجام می‌گیرد. قالب فراداده‌ای مارک، طرح فراداده‌ای هسته دوبلین، و طرح فراداده‌ای توصیف شئ (مودس) مهم‌ترین و پرکاربردترین طرح‌های فراداده‌ای توصیفی محسوب می‌شوند.

از سوی دیگر، موتورهای کاوش در زمره پراستفاده‌ترین ابزارهای جست‌وجو و بازیابی اطلاعات وب هستند. سهولت جست‌وجو در محیط آن‌ها، و قابلیت‌های جست‌وجوی متنوعی که برای سطوح مختلف کاربران فراهم نموده‌اند، بر اقبال به آن‌ها افزوده است. از این رو، دسترس‌پذیر نمودن اشیای محتوایی به ویژه پیشینه‌های فراداده‌ای از طریق موتورهای کاوش همواره مورد توجه طراحان و تولیدکنندگان آن‌ها بوده است. داده‌های کتاب‌شناختی به صورت ساختارمند در پیشینه‌های فراداده‌ای ذخیره می‌شوند. این ویژگی، به منظور پشتیبانی از امکان جست‌وجوهای مبتنی بر عناصر کتابشناختی[1]، و مدیریت کارآمد فراداده‌ها طراحی شده است. این در حالی است که اگر چه امکان جست‌وجوهای مبتنی بر بافت[2] در نظام بازیابی موتورهای کاوش امکان‌پذیر است اما، این جست‌وجوها به بخش‌های معدودی از شیء محتوایی (عنوان، متن، و ...) محدود می‌شوند.

نتایج پژوهش‌های پیشین حاکی از آن است که موتورهای کاوش وب در هنگام رویارویی با پیشینه‌های فراداده‌ای که هم اکنون از بستر نحوی زبان نشانه‌گذاری گسترش پذیر (XML) استفاده می‌کنند، رویکرد حذف برچسب را برگزیده‌اند (Taheri et al., 2009; Agha Abedi, 2010; Tabatabaie Amiri et al., 2012; Luk, Chan, Dillon, Leong, 2000; Luk, Leong, Dillon, Chan, Croft, Allan, 2002). برای نمونه، موتورهای کاوش ارزش‌های عنصر عنوان (<title>مدیریت اطلاعات</title>) به مثابه یکی از عناصر فراداده‌ای هسته دوبلین را استخراج و جست‌وجوپذیر می‌نمایند، اما قابلیت جست‌وجوی این ارزش‌ها بر اساس نام عنصر را ندارند. انتخاب رویکرد حذف برچسب از سوی موتورهای کاوش، به دلیل گسترش‌پذیر بودن زبان ایکس. ام. ال. است. از سوی دیگر برخی پژوهش‌ها، و

---

1. Element-based search
2. Contextual Search

بررسی‌های پژوهشگر نشان می‌دهند (از جمله بررسی نسخه "حافظه پنهان (cache)"، نرم‌افزارهای نمایه‌سازی موتورهای کاوش همه نویسه‌های یک شیء محتوایی دسترس‌پذیر از طریق وب (از جمله نام‌های برچسب) را استخراج می‌کنند. اما امکان بازیابی آن‌ها را فراهم نمی‌نمایند (Search Engine Watch, 2007). اکنون پرسشی که در اینجا مطرح می‌گردد آن است که چگونه می‌توان رویکرد موتورهای کاوش را در نمایه‌سازی و پیدانمایی[1] نام‌های برچسب عناصر فراداده‌ای مبتنی بر زبان نشانه‌گذاری گسترش‌پذیر تغییر داد؟

با توجه به آن که از یک سو، رویکرد کلی موتورهای کاوش در نمایه‌سازی اشیای محتوایی مبتنی بر ایکس. ام. ال.، حذف برچسب است، و از سوی دیگر، شواهد نشان می‌دهند نرم‌افزارهای نمایه‌سازی همه نویسه‌ها و اجزای اشیای محتوایی را نمایه می‌کنند، آیا می‌توان با استفاده از روش‌هایی خاص در پیاده‌سازی پیشینه‌های فراداده‌ای، با حفظ قابلیت‌های ایکس. ام. ال.، نمایه‌پذیری و پیدانمایی نام‌های برچسب عناصر آن‌ها را امکان‌پذیر نمود؟ آیا موتورهای کاوش امکان ارائه پیشینه‌های فراداده‌ای مرتبط در برابر درخواست‌های دارای نام‌های برچسب را ارائه می‌دهند؟ پژوهش حاضر به منظور پاسخگویی به این مسئله طرح‌ریزی شده است[۲].در همین راستا، و به منظور پاسخ به این مسئله، بررسی فرضیه ذیل مد نظر است: جاسازی پیشینه‌های فراداده‌ای مبتنی بر ایکس. ام. ال. در برچسب <pre> زبان نشانه‌گذاری فرامتن بر اساس روش جزیره داده‌ها[۳]، امکان نمایه‌سازی و پیدانمایی نام‌های برچسب عناصر فراداده‌ای توسط موتورهای کاوش را فراهم می‌آورد.

مرور تحلیلی پیشینه‌ها

هدف از طراحی و توسعه نظام‌های فراداده‌ای بهبود دسترسی به اشیای محتوایی است. نیل به این هدف که بر قابلیت‌های فراداده مبتنی است، در تعامل با دیگر نظام‌هایی که در همین راستا تلاش می‌کنند، قوت می‌یابد. گرایش برخی پژوهش‌های حوزه فراداده به بررسی میانکنش‌پذیری نظام‌های فراداده‌ای و موتورهای کاوش، بیانگر اهمیت تعامل یاد شده است. در یک نگاه کلی، دو جنبه اصلی در این پژوهش‌ها مد نظر قرارگرفته‌اند. نخست، پیشینه‌ها و عناصر فراداده‌ای مورد مطالعه؛ و دیگر، بستر نحوی پیاده‌سازی

---

[1] Visibility

۲. مقاله حاضر به عنوان بخشی از گزارش پژوهشی در قالب رساله دکتراست که یکی از دو فرضیه پژوهش یاد شده را پوشش می دهد. نتایج بررسی و آزمایش فرضیه دوم در مقاله دیگری ارائه شده است.

[3] Data Island Method

پیشینه‌ها و عناصر فراداده‌ای. پژوهش‌های نخستین این حوزه بر روی نمایه‌پذیری فرابرچسب‌های زبان نشانه‌گذاری فرامتن، و بازیابی اشیای محتوایی حاوی آنها تمرکز داشتند. همان طور که در منابع راهنمای برخی موتورهای کاوش نیز بیان شده است (Google, 2012; Yahoo, 2012)، نرم‌افزارهای خزنده – نمایه‌ساز موتورهای کاوش به فرا برچسب‌های یادشده واکنش مثبت نشان داده، ارزش‌های آن‌ها را نمایه‌سازی کرده‌اند. این واکنش مثبت زمینه را برای امکان بازیابی اشیای محتوایی بر اساس ارزش‌های فرابرچسب‌های نمایه شده از طریق موتورهای کاوش فراهم نموده است. نتایج پژوهش‌های ترنر و برکبیل (Turner & Brackbill, 1998)، کوئودو–تورررو (Quevedo-Torrero, 2004)، و ژانگ و دیمیتروف (Zhang & Dimitroff, 2005a) مؤید این مطلب است. در گروه دیگر پژوهش‌ها، کارآمدی عناصر مبتنی بر طرح‌های فراداده‌ای استاندارد در افزایش قابلیت بازیابی و بهبود رتبه اشیای محتوایی در میان نتایج موتورهای کاوش مورد بررسی قرار گرفت. دلیل پیدایش این گرایش، قابلیت‌ها و کارکردپذیری عناصر طرح‌های فراداده‌ای استاندارد به برچسب‌ها و فرابرچسب‌های اچ. تی. ام. ال. بود. سوکویتن (Sokvitne, 2000) و صفری (Safari, 2005) به مطالعه واکنش موتورهای کاوش به عناصر طرح فراداده‌ای هسته دوبلین در بستر اچ. تی. ام. ال. پرداختند. نتایج هر دو پژوهش بر عدم و یا کاهش نمایه‌پذیری ارزش‌های عناصر فراداده‌ای طرح‌های استاندارد توسط موتورهای کاوش، و به پیروی از آن عدم یا کاهش قابلیت بازیابی اشیای محتوایی حاوی آن عناصر تأکید نمودند.

گروه دو دیگر پژوهش‌های حوزه میانکنش‌پذیری، به اثربخشی برچسب‌ها و فرابرچسب‌های زبان نشانه‌گذاری فرامتن و عناصر فراداده‌ای طرح‌های استاندارد به صورت ترکیبی و با رویکرد مقایسه‌ای اختصاص داشت. واکنش مثبت موتورهای کاوش به فرابرچسب‌های زبان نشانه‌گذاری فرامتن، و واکنش منفی آنها به نمایه‌سازی عناصر فراداده‌ای طرح‌های استاندارد، موجب پیدایش چنین رویکردی در پژوهش‌های حوزه میانکنش‌پذیری بود. یافته‌های پژوهش‌های هنشاو و والاسکاس (Henshaw & Valauskas, 2001)، ژانگ و دیمیتروف (Zhang and Dimitroff, 2004)، ژانگ و دیمیتروف (Zhang and Dimitroff, 2005b)، محمد (Mohamed, 2006)، و شریف (Sharif, 2007) تأثیر بهتر فرا برچسب‌های زبان نشانه‌گذاری فرامتن در دسترس‌پذیر نمودن اشیای محتوایی وب از طریق موتورهای کاوش را نسبت به عناصر فراداده‌ای طرح‌های استاندارد نشان می‌دادند. تمرکز صرف بر روی قالب زبان نشانه‌گذاری فرامتن در پژوهش‌های پیشین، مبدأ پیدایش پژوهش‌هایی جدید شد. مهم‌ترین ویژگی آخرین گروه پژوهش، تغییر بستر نحوی پیشینه‌های فراداده‌ای به زبان نشانه‌گذاری گسترش‌پذیر بود. پژوهش‌های

این گروه، راه حل افزایش میانکنش پذیری موتورهای کاوش و نظام‌های فراداده‌ای را در تغییر بستر نحوی به زبان نشانه‌گذاری گسترش پذیر یافتند. یافته‌های پژوهش‌های مذکور حاکی از آن بود که انتخاب بستر نحوی ایکس. ام. ال. موجب افزایش نمایه‌پذیری و قابلیت بازیابی ارزش‌های عناصر فراداده‌ای توسط موتورهای کاوش می‌گردد (Kelly, 2006; Taheri et al., 2009; Agha Abedi, 2010; Tabatabaie Amiri, et al., 2012). پژوهش‌های حوزه میانکنش‌پذیری دارای نقاط مشترک ویژه‌ای بودند. همه آن‌ها بر اساس روش تجربی انجام شدند. به استثنای پژوهش آقا عابدی (Agha Abedi, 2010)، تمامی پژوهش‌ها رویکرد جاسازی را رویکرد استفاده از فراداده انتخاب نمودند. یافته‌های پژوهش‌های این حوزه بیانگر آن بود که فرا برچسب‌های زبان نشانه‌گذاری فرامتن برای موتورهای کاوش شناخته‌شده هستند، و ارزش‌های آن‌ها توسط نرم‌افزارهای خزنده–نمایه ساز کاملاً نمایه‌سازی شده، و موجب دسترس‌پذیری اشیای محتوایی مرتبط می‌شوند.

همچنین فرا برچسب‌های اچ. تی. ام. ال. رتبه اشیای محتوایی را در فهرست نتایج موتورهای کاوش بهبود می‌بخشند (Turner & Brackbill, 1998; Quevedo-Torrero, 2004; Zhang & Dimitroff, 2005a). اما این موتورها به ارزش‌های عناصر فراداده‌ای طرح‌های استاندارد که در بستر نحوی اچ. تی. ام. ال. پیاده‌سازی شده‌اند، و نیز نام‌های برچسب عناصر واکنش مثبتی نشان ندادند (Sokvitne, 2000; Safari, 2005). خلاء پژوهشی که در پژوهش‌های حوزه میانکنش پذیری مشاهده می‌شود، عدم توجه به نمایه‌سازی و پیدانمایی نام‌های برچسب عناصر به عنوان اصطلاحات نمایه‌ای توسط موتورهای کاوش است (Hariri, et al., 2013). بدیهی است اگر چنین قابلیتی برای عناصر فراداده‌ای که به توصیف و کشف اشیای محتوایی می‌پردازند، به وجود آید، کاربران نهایی می‌توانند دامنه جست‌وجوهای خود را به ارزش‌های عناصری خاص محدود نمایند، و همانند پایگاه‌های اطلاعاتی تمام متن و فهرست‌های رایانه‌ای که امکان جست‌وجوهای مبتنی بر فیلد ارائه می‌دهند، در موتورهای کاوش نیز امکان جست‌وجوهای مبتنی بر عنصر میسر می‌گردد.

روش پژوهش

پژوهش حاضر از نظر نوع یک پژوهش کاربردی است، و روش پژوهش، روش تجربی (از دو گروه گواه و آزمون بهره می‌برد)، و مبتنی بر فناوری(در بستر فناوری‌های اطلاعاتی انجام‌شده و به مطالعه آن‌ها می‌پردازد) است. جامعه پژوهش را ۶۰۰ پیشینه فراداده‌ای در قالب گروه گواه، کاملاً مبتنی بر فرانماهای استانداردهای فراداده‌ای هسته دوبلین، مارک ۲۱، و طرح فراداده‌ای توصیف شیء (مودس)،

وگروه آزمون، مبتنی بر فرانماهای یادشده، اما تولیدشده بر مبنای روش جزیره داده‌ها، تشکیل می‌دادند. گروه گواه شامل ۳۰۰ پیشینه (برای هر یک از طرح‌های فراداده‌ای مورد مطالعه ۱۰۰ پیشینه)، و گروه آزمون دربردارنده ۳۰۰ پیشینه (همانند گروه گواه) بود. گروه‌های یادشده بر روی دو وب‌سایت مستقل با شرایط کاملا یکسان منتشر (بدون هیچ پیوندی از، یا به دو وب‌سایت)، و با استفاده از پیشرفته‌ترین روش-های ارائه‌شده (نقشه‌های سایت مبتنی بر اچ. تی. ام. ال.، مبتنی بر متن، منبعِ منبع[1]، و مبتنی بر ایکس. ام. ال.) به موتورهای کاوش گوگل و یاهو معرفی شدند. یک وب سایت به گروه گواه (http://www.tagnamemeta-cont.ir) و یک وب سایت به گروه آزمون (http://www.tagnamemeta-ex1.ir) اختصاص یافت. وب سایت‌های یاد شده از تاریخ دی ماه ۱۳۹۰ به مدت یک سال بر روی وب دسترس پذیر هستند. طرح‌های فراداده‌ای مورد مطالعه به این دلیل انتخاب شدند که هر کدام دارای ویژگی خاصی از لحاظ ساختاری بودند و ممکن بود بر نمایه‌سازی و پیدانمایی آن‌ها توسط موتورهای کاوش تأثیر بگذارد.

هر یک از طرح‌های فراداده‌ای برگزیده دارای فرانماهای مبتنی بر زبان نشانه‌گذاری گسترش‌پذیر هستند. ۳۰۰ پیشینه فراداده‌ای مربوط به گروه گواه کاملاً بر اساس این فرانماها ساخته شده، و با استفاده از ابزارهای خاص (Styles Studio و MARCXML MARC Validation Stylesheets) اعتبارشان سنجیده شد. طراحی و تعیین این ساختار برای گروه گواه، بر اساس نتایج پژوهش‌های لوک و دیگران (Luk et al., 2000) ،طاهری و همکاران (Taheri, et al., 2009) ، آقاعابدی (Agha Abedi, 2010) ، و طباطبایی امیری و همکاران (Tabatabaie Amiri, et al., 2012) بود، که نشان دادند رویکرد موتورهای کاوش برای نمایه‌سازی اشیای محتوایی مبتنی بر ایکس. ام. ال.، رویکرد "حذف برچسب" است. ۳۰۰ پیشینه بعدی مربوط به هر طرح فراداده‌ای در گروه آزمون بر اساس روش جزیره داده‌ها طراحی و ایجاد شدند. بر اساس این روش، کل محتوای هر پیشینه مبتنی بر ایکس. ام. ال. در برچسب <pre> درون برچسب <body> یک مدرک مبتنی بر زبان نشانه‌گذاری فرامتن جاسازی شد، و نشانه‌های آغاز برچسب "<" و پایان برچسب ">" به نشانه‌های "رها از موجودیت" (به ترتیب"<" و ">") تبدیل شدند. این پیشینه‌ها نیز ویژگی‌های قالب پیشین خود را کاملاً حفظ می‌کردند. نام برچسب <pre> که خلاصه "متن از پیش قالب‌بندی شده"[2] است، بدین معناست که ارزشی که در آن قرار می‌گیرد، مربوط به شیئی محتوایی است که دارای قالب و ساختار متنی ویژه است، و با قرار گرفتن آن در برچسب <pre>، قالب قبلی خود

[1] Resource of Resource
[2] Pre-formatted text

(ساختار، نشانه‌ها، و روابط معناشناختی عناصر) را حفظ خواهد نمود. بدین ترتیب می‌توان علاوه بر بهره‌مندی از قابلیت‌های قالب قبلی، از ویژگی‌های زبان نشانه‌گذاری فرا متن (نمایش، طراحی، چندرسانه‌ای، و ...) نیز استفاده کرد (Kyrnin, 2001; Woychowsky, 2003; Expertrating, 2012; Microsoft Developers Network, 2012; www.w3schools.com).پیشینه‌های فراداده‌ای هسته دوبلین و مارک ۲۱ از وب سایت http://www.archive.org و از مجموعه اشیای محتوایی کتابخانه دیجیتالی کالیفرنیا، و از یک رده موضوعی (Authors) مورد علاقه پژوهشگر، با داشتن یکصد شیء محتوایی، و بدون تأثیر بر زمینه و نتایج پژوهش انتخاب گردیدند. بخش عمده پیشینه‌های فراداده‌ای مربوط به طرح فراداده‌ای توصیف شیء (مودس) از وب سایت کتابخانه کنگره (http://lccn.loc.gov)، و بخشی هم از طریق تبدیل پیشینه‌های فراداده‌ای مارک ۲۱ در بستر ایکس. ام. ال. با استفاده از ابزار "MARCXML Toolkit" تهیه شدند. گام بعدی تعیین موتورهای کاوش وب برای بررسی واکنش آن‌ها به پیشینه‌های فراداده‌ای مورد مطالعه بود. چنان چه بسیاری از منابع معتبر بیان می‌دارند، موتورهای کاوش گوگل و یاهو[1]در بین موتورهای کاوش وب دارای بالاترین کیفیت خدمات، و بالاترین بسامد استفاده هستند (Alexa, 2012; Lewis, 2012;Campex, 2012;ComScore, 2012). از این رو در این پژوهش، دو موتور اشاره‌شده به عنوان نماینده موتورهای کاوش وب برگزیده شدند.

روش گردآوری داده‌ها در این پژوهش از طریق مشاهده طراحی شده بود. ابزار گردآوری داده-ها، سیاهه وارسی (چک لیست) طراحی‌شده به عنوان ابزار ثبت مکانیکی با استفاده از نرم‌افزار آماری SPSS توسط پژوهشگر بود. همچنین در مواردی از عکس‌های فوری، و ذخیره آن‌ها از طریق نرم‌افزار پینت، به عنوان ابزار الکترونیکی استفاده شد. سیاهه وارسی قضاوت در اختیار پنج نفر از صاحب‌نظران باتجربه حوزه ذخیره و بازیابی اطلاعات قرار داده شد، و روایی آن تأیید شد. لازم به ذکر است از آنجا که ابزار گردآوری داده‌ها سیاهه وارسی بود، نیاز به سنجش پایایی نداشت.گردآوری داده‌ها از اواسط بهمن‌ماه ۱۳۹۰ تا پایان اسفندماه ۱۳۹۰ صورت گرفت.

۱. از جولای ۲۰۱۱ پشتیبانی از بخش معرفی مستقیم وب سایت‌ها به یاهو، و از اواسط آگوست ۲۰۱۱ ابزار مدیریت وب سایت (Web Master Tools) آن به موتور کاوش بینگ واگذار شده است. همچنین از اوایل سال ۲۰۱۲، محتوای وب سایت‌هایی که از روش‌های معرفی مستقیم یا ابزار مدیران وب‌سایت‌های یاهو استفاده نموده بودند، مجدداً توسط نرم افزار خزنده-نمایه‌ساز بینگ (msnbot) مورد پیمایش و نمایه سازی قرار گرفته‌اند. این مطلب در مدارک یاهو پیش‌تر توصیه شده بود.

تجزیه و تحلیل داده‌ها

برای آزمون فرضیه بر اساس آمار استنباطی، و به منظور تعیین تفاوت وضعیت نمایه‌سازی و پیدانمایی نام‌های برچسب عناصر پیشینه‌های فراداده‌ای گروه آزمون و گروه گواه از آزمون آماری "مقایسه زوجی[1]" استفاده شد. دلیل استفاده از این آزمون، انجام پژوهش به روش تجربی که دارای گروه‌های آزمون و گواه است، و نوع یافته‌های به دست آمده بود.

جدول شماره ۱. مقایسه زوجی بین نمایه‌پذیری و پیدانمایی نام‌های برچسب عناصر فراداده‌ای هسته دوبلین، مارک ۲۱، و طرح فراداده‌ای توصیف شیء عضو گروه آزمون و گروه گواه در محیط موتورهای کاوش گوگل و یاهو[2]

گروه‌ها	موتورهای کاوش و طرح‌های فراداده‌ای مورد بررسی	آزمون =۱؛ گواه= ۲	رسته[3]	پیشینه	نسبت مشاهده‌شده[4]	نسبت آزمون[5]	معنی‌داری مجانبی (۲ دمی)[6]
گواه و آزمون	موتور کاوش گوگل – طرح فراداده‌ای هسته دوبلین	گروه ۱	۱	۱۰۰	۱,۰۰	.۵۰	.۰۰۰ (a)
		گروه ۲	۲	۱۰۰	.۰۰	.۵۰	
		مجموع		۲۰۰	۱,۰۰	۱,۰۰	
	موتور کاوش یاهو – طرح فراداده ای هسته دوبلین	گروه ۱	۱	۱۰۰	۱,۰۰	.۵۰	.۰۰۰ (a)
		گروه ۲	۲	۱۰۰	.۰۰	.۵۰	
		مجموع		۲۰۰	۱,۰۰	۱,۰۰	
	موتور کاوش گوگل- قالب فراداده‌ای مارک۲۱	گروه ۱	۱	۱۰۰	۱,۰۰	.۵۰	.۰۰۰ (a)
		گروه ۲	۲	۱۰۰	.۰۰	.۵۰	
		مجموع		۲۰۰	۱,۰۰	۱,۰۰	
	موتور کاوش یاهو – قالب فراداده‌ای مارک۲۱	گروه ۱	۱	۱۰۰	۱,۰۰	.۵۰	.۰۰۰ (a)
		گروه ۲	۲	۱۰۰	.۰۰	.۵۰	
		مجموع		۲۰۰	۱,۰۰	۱,۰۰	
	موتور کاوش گوگل –	گروه ۱	۱	۱۰۰	۱,۰۰	.۵۰	.۰۰۰ (a)

---

[1] Binomial Test

۲. بر مبنای تقریب Z و P<0/001

[3] Category

[4] Observed proportion

[5] Test proportion

[6] Asymptotic significance 2-tailed

	طرح فراداده‌ای توصیف	گروه ۲	۲	۱۰۰	.۰۰	.۵۰	
	شیء	مجموع		۲۰۰	۱٫۰۰	۱٫۰۰	
	موتور کاوش یاهو –	گروه ۱	۱	۱۰۰	۱٫۰۰	.۵۰	.۰۰۰ (a)
	طرح فراداده‌ای توصیف	گروه ۲	۲	۱۰۰	.00	.۵۰	
	شیء	مجموع		۲۰۰	۱٫۰۰	۱٫۰۰	

داده‌های جدول ۱ فرضیه پژوهش را تأیید کرده، و بیانگر تفاوت معنادار میان وضعیت نمایه‌سازی و پیدانمایی نام‌های برچسب عناصر فراداده‌ای مبتنی بر طرح‌های فراداده‌ای مورد مطالعه توسط موتورهای کاوش گوگل و یاهو در گروه آزمون با گروه گواه می‌باشد. از این رو، جاسازی پیشینه‌های فراداده‌ای مبتنی بر زبان نشانه‌گذاری گسترش‌پذیر در برچسب <pre> زبان نشانه‌گذاری فرامتن (مبتنی بر روش جزیره داده‌ها) باعث افزایش نمایه‌پذیری و پیدانمایی نام‌های برچسب عناصر پیشینه‌های فراداده‌ای گردید.

بحث و نتیجه‌گیری

نمایه‌سازی و پیدانمایی نام‌های برچسب عناصر فراداده‌ای توسط موتورهای کاوش وب، به عنوان پراستفاده‌ترین ابزار کاوش وب، موجب بهبود بازیابی و در نتیجه افزایش دسترس‌پذیری به پیشینه‌های فراداده‌ای، و به پیروی از آن اشیای محتوایی وب می‌شود. بر اساس نتیجه آزمون انجام‌شده برای فرضیه پژوهش، استفاده از برچسب <pre> برای جاسازی پیشینه‌های فراداده‌ای مبتنی بر ایکس. ام. ال. بر اساس "روش جزیره داده‌ها"، موجب نمایه‌سازی و پیدانمایی نام‌های برچسب عناصر پیشینه‌های فراداده‌ای، و حتی خصایص[1](مانند شماره فیلد و نشانگرهای آن‌ها در پیشینه‌های مارک ۲۱) برچسب‌ها که بر بهبود بازیابی و افزایش مانعیت و ربط تأثیر فراوانی دارند، توسط موتورهای کاوش وب می‌گردد. مرورگرهای پرکاربرد اینترنت (اکسپلورر و فایرفاکس) نیز اشیای محتوایی دارای برچسب <pre> را به خوبی پشتیبانی می‌کنند، و هنگام نمایش آن‌ها، قابلیت‌های قالب قبلی شیء را به صورت معنادار اعمال می‌نمایند.این شیوه برخورد با نام‌های برچسب عناصر توسط موتورهای کاوش، بهترین رویکردی است که در منابع مختلف برای نمایه‌سازی اشیای محتوایی مبتنی بر زبان نشانه‌گذاری گسترش‌پذیر بیان شده است(Luk et al., 2000; Qin, 2000; Luk et al., 2002; Gigee, 2006; Gill, 2008; Taheri and Hariri, 2012).

---
[1] Attributes

از سوی دیگر، نمایه‌سازی و پیدانمایی نام‌های برچسب به عنوان اصطلاحات نمایه‌ای توسط موتورهای کاوش موجب افزایش دسترس‌پذیری به پیشینه‌های فراداده‌ای و اشیای محتوایی شده، این موتورها را به دروازه‌های اطلاعاتی تبدیل می‌نماید که کیفیت محتوای جست‌وجوپذیر را تضمین می‌نمایند. زیرا محتوای استخراج‌شده از معتبرترین نظام‌های اطلاعاتی وب مربوط به بافت اطلاعات فرهنگی (کتابخانه‌ای، آرشیوی، و موزه‌ای) و از درون پیشینه‌های فراداده‌ای خواهد بود. همچنین امکان جست‌وجوهای مبتنی بر عنصر، باعث افزایش مانعیت در بازیابی پیشینه‌های فراداده‌ای در محیط موتورهای کاوش می‌شود. بنابراین طراحی و تولید پیشینه‌های فراداده‌ای بر اساس روش معرفی‌شده در این پژوهش، به تولید ارزش افزوده برای نظام‌های فراداده‌ای منجر خواهد شد. اگرچه در این پژوهش فقط دو موتور کاوش مورد بررسی قرار گرفتند، اما بررسی‌ها نشان می‌دهند بیشتر موتورهای کاوش وب فاقد نرم‌افزار روبات هستند، و برای تأمین اطلاعات مورد نیاز کاربران خود از نتایج موتورهای کاوش گوگل و یاهو استفاده می‌کنند.

ذکر این مطلب در مورد یافته‌های پژوهش ضروری است که با مطالعه بر روی سه طرح فراداده‌ای معتبر با ساختارهای مختلف (درختی خانوادگی و درختی ساده)، و همخوان و ناهمخوان با زبان طبیعی، و انتشار ۱۰۰ پیشینه مبتنی بر هر طرح و بر اساس روش مورد استفاده در پژوهش حاضر، و نتایج به دست آمده از آزمون فرضیه، می‌توان چنین نتیجه گرفت که امکان تعمیم‌یافته‌ها در حوزه میانکنش‌پذیری نظام‌های فراداده‌ای و موتورهای کاوش وجود دارد. مقایسه یافته‌های پژوهش حاضر با پژوهش‌های پیشین بیانگر این مطلب است که روش استفاده‌شده در تولید پیشینه‌های فراداده‌ای گروه آزمون، گامی رو به جلو در دسترس‌پذیری پیشینه‌های فراداده‌ای از طریق موتورهای کاوش وب محسوب می‌گردد. در پژوهش‌هایی که زبان نشانه‌گذاری فرامتن بستر نحوی پیاده‌سازی پیشینه‌های فراداده‌ای انتخاب شده بود، نمایه‌پذیری و پیدانمایی عناصر فراداده‌ای توسط موتورهای کاوش سطح پایینی را تجربه می‌نمود.

موتورهای کاوش فقط ارزش فرابرچسب‌های توصیفی اچ. تی. ام. ال. به عنوان عناصر فراداده‌ای را نمایه‌سازی می‌نمودند، و نسبت به ارزش‌های عناصر دیگر طرح‌های فراداده‌ای از جمله هسته دوبلین بی‌تفاوت عمل می‌کردند. این عدم واکنش نسبت به نام‌های برچسب عناصر طرح‌های فراداده‌ای استاندارد، و حتی فرابرچسب‌های اچ. تی. ام. ال. شدت بیشتری می‌یافت (شریف، ۱۳۸۶؛ Turner and Brackbill, 1998; Sokvitne, 2000; Henshaw and Valauskas, 2001; Quevedo-Torrero, 2004; Zhang and Dimitroff; 2004; Zhang and Dimitroff, 2005a; Zhang and Dimitroff, 2005b; Safari, 2005; Mohamed, 2006). بر اساس نتایج پژوهش‌های طاهری و همکاران (Taheri et al., 2009)، آقاعابدی (Agha Abedi, 2010)، و طباطبایی

امیری و همکاران (Tabatabaie Amiri et al., 2012) امکان نمایه‌پذیری و پیدانمایی نام‌های برچسب عناصر برای پیشینه‌های فراداده‌ای که در قالب ایکس. ام. ال. پیاده‌سازی شده بودند نیز وجود نداشت، و نرم‌افزارهای خزنده- نمایه‌ساز موتورهای کاوش وب صرفاً به نمایه‌سازی ارزش‌های عناصر آن‌ها بسنده نموده بودند، و یا این که نام‌های برچسب مرتبط را در نتایج جست‌وجو پدیدار نمی‌نمودند. در این بخش از مقاله، با توجه به تأثیر صد درصدی روش مورد استفاده در این پژوهش برای تولید پیشینه‌های فراداده‌ای با نام‌های برچسب نمایه‌پذیر و قابل پیدانمایی که موجب بهبود بازیابی اشیای محتوایی در محیط موتورهای کاوش وب می‌شوند، و به منظور بهره‌مندی از مزایای آن توسط مخاطبان پژوهش، یعنی تولیدکنندگان پیشینه‌های فراداده‌ای و کاربران نهایی، چهار الگو[1] در قالب پیشینه‌های فراداده‌ای طراحی‌شده‌اند که در ذیل ارائه می‌گردند. دو الگوی نخست برای تولیدکنندگان پیشینه‌های فراداده‌ای، و دو الگوی دیگر برای کاربران نهایی هستند.

الگوی نخست. نمونه‌ای از یک پیشینه مبتنی بر هسته دوبلین بر پایه روش جزیره داده‌ها دارای ارزش عناصر به زبان فارسی،جاسازی‌شده در برچسب <pre> و با پسوند فایل ".html"[2]

۱. لازم به ذکر است الگوهای ارائه شده در این مقاله در فرایند ثبت به عنوان اختراع ملی قرار گرفته اند، و هر گونه نسخه برداری، اقتباس، و استفاده از این الگوها بر خلاف قوانین جمهوری اسلامی ایران محسوب شده، از سوی مراجع قضایی پیگرد قانونی دارد.

۲. پیشینه‌های الگوی مبتنی بر قالب فراداده‌ای مارک ۲۱ و طرح فراداده‌ای توصیف شیء که در اصل پژوهش ارائه شده‌اند نیز شبیه الگوی نخست هستند که به منظور جلوگیری از افزایش حجم مقاله ارائه نشده‌اند.

```
<html><head>
<meta http-equiv="content-type" content="text/html; charset=UTF-8">
    <title>DCXML</title>
    <style type="text/css" id="internalStyle">>
        body { color: black; background: white; }
        code,pre { font-family: "Courier New", Courier, monospace; font-size: 100%; }
        em { color: red; }
        table { background: #CCCCCC; }
        th { font-weight: bold; font-size: 120%; }
        table.header { background: white; }
    </style>
</head>
<body>
<h2>DCXML</h2>
<table summary="Static Repository Example" width="80%" border="2">
  <tbody><tr>
    <td>
<pre>
&lt;?xml version="1.0"?&gt;<br>
&lt;metadata
  xmlns="http://www.ukoln.ac.uk/metadata/dcdot/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.ukoln.ac.uk/metadata/dcdot/ http://www.ukoln.ac.uk/metadata/dcdot/dcdot.xsd"
  xmlns:dc="http://purl.org/dc/elements/1.1/"&gt;
  &lt;dc:title xml:lang="فارسي" &gt;
    پژوهشگاه علوم و فرهنگ اسلامي
  &lt;/dc:title&gt;
  &lt;dc:creator&gt;
    مدير وب سايت
  &lt;/dc:creator&gt;
  &lt;dc:subject xml:lang="فارسي" &gt;
    دانشگاه باقرالعلوم; سمتا;
    تبليغات; پاسخگويي به مسائل شرعي;
    پژوهشگاه علوم و فرهنگ اسلامي; ;
    دفتر تبليغات اسلامي حوزه علميه
    قم; پورتال نشريات; توسعه زير
    ساختهاي فناوري اطلاعات سيگما
  &lt;/dc:subject&gt;
  &lt;dc:description xml:lang="فارسي" &gt;
    وب سايت سازماني پژوهشگاه علوم و
    فرهنگ اسلامي وابسته به دفتر
    تبليغات اسلامي حوزه علميه قم
  &lt;/dc:description&gt;
  &lt;dc:publisher&gt;
    واحد نشر الكترونيكي پژوهشگاه
  &lt;/dc:publisher&gt;
  &lt;dc:contributor&gt;
    با همكاري علي حجازي
  &lt;/dc:contributor&gt;
  &lt;dc:date&gt;
    1391/01/31
  &lt;/dc:date&gt;
  &lt;dc:type&gt;
    Text
  &lt;/dc:type&gt;
  &lt;dc:format&gt;
    text/html; charset=utf-8
  &lt;/dc:format&gt;
  &lt;dc:format&gt;
    204596 bytes
  &lt;/dc:format&gt;
  &lt;dc:identifier&gt;
    http://www.isca.ac.ir
  &lt;/dc:identifier&gt;
  &lt;dc:language&gt;
    فارسي-عربي-انگليسي
  &lt;/dc:language&gt;
  &lt;dc:relation&gt;
    http://www.islamicdoc.org
  &lt;/dc:relation&gt;
  &lt;dc:coverage&gt;
    گزارش فعاليت هاي پژوهشگاه
  &lt;/dc:coverage&gt;
  &lt;dc:rights&gt;
    پژوهشگاه علوم و فرهنگ اسلامي
  &lt;/dc:rights&gt;
&lt;/metadata&gt;
</pre>
    </td>
  </tr>
</tbody></table>
</body></html>
```

چنان‌که پیش‌تر گفته شد، مرورگرهای وب (از جمله اینترنت اکسپلورر و فایرفاکس) قادر به نمایش نویسه‌های رها از موجودیت به شکل طبیعی هستند، و پیشینه‌های تولیدشده از لحاظ نمایشی تفاوتی

با دیگر پیشینه‌ها ندارند. الگوی ارائه‌شده بر اساس روش "جزیره داده‌ها"، در اصل پیشینه فراداده‌ای مبتنی بر ایکس. ام. ال. است که درون برچسب <pre> جاسازی‌شده، و پسوند فایل آن به "html." تغییر یافت. به دلیل قابلیت‌های برچسب <pre> که پیش‌تر به آن‌ها اشاره شد، پیشینه‌های جاسازی‌شده، ساختار و ویژگی‌های بستر نحوی زبان نشانه‌گذاری گسترش‌پذیر را حفظ نموده‌اند، و از قابلیت‌های قالب جدید (زبان نشانه‌گذاری فرامتن) نیز سود می‌برند.

اما مهم‌ترین ویژگی که در پیشینه نمونه این الگو وجود دارد، استفاده از نشانه‌های "<" و ">" است که به ترتیب برای نشانه‌های آغاز برچسب "<" و پایان برچسب ">" برای نام‌های برچسب در هنگام طراحی پیشینه‌هاست. این نشانه‌ها باعث می‌گردند موتور کاوش در هنگام برخورد با برچسب‌ها، آن‌ها را برچسب تلقی نکند، و به عنوان اصطلاحات نمایه‌ای، نمایه‌سازی نموده، و در فهرست نتایج خود پدیدار نماید. نکته جالب‌توجه آن است که بر خلاف موتورهای کاوش، مرورگرهای وب این نشانه‌ها را به خوبی می‌شناسند، و هنگام نمایش به نشانه‌های آغاز برچسب "<" و پایان برچسب ">" تبدیل می‌کنند. این نویسه‌ها که به نویسه‌های "رها از موجودیت" مشهور هستند، برای تمامی نرم‌افزارهای پردازشگر اشیای محتوایی مبتنی بر ایکس. ام. ال. نیز شناخته‌شده بوده، و با آن‌ها به عنوان نشانه‌های "<" و ">" برخورد می‌شود. نمونه برچسب‌های عنصر عنوان هسته دوبلین، <dc:title></dc:title> که شکل نمایشی آن‌ها در مرورگرها بدین صورت خواهد بود: <dc:title></dc:title>

اگر چه نمایه‌سازی نام‌های برچسب عناصر توسط موتورهای کاوش، امکان انجام جست‌وجوهای مبتنی بر عنصر را در محیط آن‌ها ایجاد می‌کند، و مانعیت فرایند بازیابی را افزایش می‌دهد، اما به علت استفاده محدود موتورهای کاوش از نظام‌های معنایی، و نیز عدم استفاده از فهرست‌های مستند نام‌ها از یک سو، و صرفاً ثبت شکل پذیرفته‌شده نام‌ها و اصطلاحات در پیشینه‌های فراداده‌ای کتاب‌شناختی از سوی دیگر، قابلیت بازیابی پیشینه‌های مورد نظر کاربران از طریق دیگر شکل‌های نام‌ها و اصطلاحات وجود ندارد. بنابراین موتورهای کاوش از جامعیتی که نظام‌های معنایی برای فرایند بازیابی فراهم می‌نمایند، بهره‌مند نمی‌گردند. به منظور حل این مشکل، یک الگوی مناسب و کاربردی برای استفاده از جامعیت فراهم‌شده توسط نظام‌های معنایی در محیط موتورهای کاوش پیشنهاد می‌گردد.

بر اساس روش مورد استفاده در پژوهش، برای تولید پیشینه‌هایی با نام‌های برچسب نمایه‌پذیر و قابل ارائه در نتایج کاوش، می‌توان پیشینه‌های فراداده‌ای مستند مبتنی بر ایکس. ام. ال. را در پیشینه‌های فراداده‌ای کتاب‌شناختی جاسازی نمود.

در این صورت، پیشینه‌هایی در قالب "مجموعه"[1] تولید خواهند شد که علاوه بر نام‌های برچسب و ارزش‌های عناصر کتاب‌شناختی، نام‌های برچسب و ارزش‌های عناصر مستند مرتبط با عناصر کتاب‌شناختی آن پیشینه را در بر می‌گیرند، و این ویژگی را دارا هستند که توسط موتورهای کاوش نمایه‌سازی شده، و در نتایج جست‌وجوی آن‌ها پدیدار شوند. بنابراین، پیشینه‌های فراداده‌ای بر اساس دیگر شکل‌های نام‌ها و اصطلاحات قابل بازیابی خواهند بود. این ویژگی نیز برای حوزه میانکنش پذیری نظام‌های فراداده‌ای و موتورهای کاوش وب ارزش افزوده تولید خواهد نمود.

```
<html><head>
<meta http-equiv="content-type" content="text/html; charset=UTF-8">
    <title>DCXML</title>
    <style type="text/css" id="internalStyle">>
        body { color: black; background: white; }
        code,pre { font-family: "Courier New", Courier, monospace; font-size: 100%; }
        em { color: red; }
        table { background: #CCCCCC; }
        th { font-weight: bold; font-size: 120%; }
        table.header { background: white; }
    </style>
</head>
<body>
<h2>DCXML</h2>
<table summary="Static Repository Example" width="80%" border="2">
  <tbody><tr><td>
 <pre>
  &lt;?xml version="1.0" encoding="UTF-8"?&gt;
 &lt;mads:mads xmlns:mads="http://www.loc.gov/mads/v2" xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.loc.gov/mads/v2
  http://www.loc.gov/standards/mads/mads-2-0.xsd" version="2.0"&gt; &lt;mads:authority geographicSubdivision="indirect"&gt;
 &lt;mads:topic authority="lcsh"&gt;Child authors&lt;/mads:topic&gt;&lt;/mads:authority&gt;
 &lt;&lt;/mads:related&gt;&lt;mads:variant type="other"&gt;
 &lt;mads:topic&gt;Children as authors&lt;/mads:topic&gt;
 &lt;/mads:variant&gt;&lt;mads:variant type="other"&gt;
 &lt;mads:topic&gt;Juvenile authors&lt;/mads:topic&gt;
 &lt;/mads:variant&gt;&lt;mads:note type="source"&gt;Hennepin
 &lt;/mads:note&gt;&lt;mads:note type="source"&gt;LC database, June 26, 1995 (Juvenile authors)
 &lt;/mads:note&gt;&lt;mads:note type="source"&gt;97122054: Subject guide to Australian children's books in print, 1995
 &lt;/mads:note&gt;&lt;mads:note&gt;Here are entered works on child authors as a class of persons. Collections of works
```

[1] Collection Record

```
published for children, including literary works, nonliterary works, or both, are entered under Children's literature.
Works on books for children considered as physical objects are entered under Children's books. Works on the reading
interests of children or books or recommended for children, are entered under Children--Books and reading. Collections of
literary works or individual literary works written by children under 15 years of age are entered under Children's writings
&lt;/mads:note&gt;&lt;mads:note type="subject example"&gt;Notes under Children--Books and reading; Children's books;
Children's literature; Children's writings&lt;/mads:note&gt;
&lt;mads:identifier type="lccn"&gt;sh 85023524  &lt;/mads:identifier&gt;&lt;mads:classification&gt;BF723.A8Child psychology
&lt;/mads:classification&gt;
&lt;mads:classification&gt;
PN171.C5Literature&lt;/mads:classification&gt;&lt;mads:recordInfo&gt;  &lt;mads:recordOrigin&gt;Converted from MARCXML to
MADS version 2.0 (Revision 2.10)&lt;/mads:recordOrigin&gt;
&lt;mads:recordContentSource authority="marcorg"&gt;DLC&lt;/mads:recordContentSource&gt;
&lt;mads:recordChangeDate encoding="iso8601"&gt;20090402075850.0
&lt;/mads:recordChangeDate&gt;&lt;mads:recordIdentifier source="DLC"&gt;sh 85023524
&lt;/mads:recordIdentifier&gt; &lt;/mads:recordInfo&gt;&lt;/mads:mads&gt;
</pre></td></tr><tr><td><pre>&lt;?xml version="1.0"?&gt;&lt;dc xmlns:dc="http://purl.org/dc/elements/1.1/"&gt;
&lt;dc:title&gt;Almond blossom; a collection of verse and prose,&lt;/dc:title&gt;&lt;dc:creator&gt; Johnston, K M , &lt;
/dc:creator&gt;&lt;dc:type&gt;text&lt;/dc:type&gt;
&lt;dc:publisher&gt;London, S. Low, Marston&lt;/dc:publisher&gt;&lt;dc:date&gt;[192-?]&lt;/dc:date&gt;&lt;dc:language&gt;eng
&lt;/dc:language&gt;
&lt;dc:subject&gt;Child authors.&lt;/dc:subject&gt;
&lt;/dc&gt;</pre></td></tr></tbody></table></body></html>
```


الگوی دوم. نمونه‌ای از یک پیشینه مجموعه هسته دوبلین بر پایه روش جزیره داده‌ها، ترکیبی از پیشینه کتاب‌شناختی و پیشینه مستند مرتبط با آن جاسازی‌شده در برچسب <pre> و با پسوند فایل "html."

همان طور که در الگو ارائه‌شده مشاهده می‌شود، علاوه بر پیشینه‌های کتاب‌شناختی تولیدشده بر مبنای طرح فراداده‌ای هسته دوبلین، پیشینه مستند موضوعی مرتبط با آن که از پایگاه مستند موضوعی کتابخانه کنگره آمریکا استخراج‌شده، و بر اساس طرح فراداده‌ای توصیف مستند (مدس) ایجاد گردیده نیز در پیشینه مجموعه جاسازی شده است. روش استفاده‌شده برای تولید پیشینه‌های مجموعه از لحاظ فنی معتبر بوده، و برای موتورهای کاوش نمایه‌پذیر و قابل پیدانمایی است. بدیهی است چنین امکانی برای پیشینه‌های مستند نام‌ها و پیشینه‌های کتابشناختی مبتنی بر دیگر طرح‌های فراداده‌ای نیز وجود دارد.لازم به ذکر است به علت ویژگی‌های خاص قالب زبان نشانه‌گذاری فرامتن، پیشینه مستند در یک برچسب <pre> متفاوت با برچسب پیشینه کتابشناختی، در پیشینه مجموعه جاسازی شده است، و تمامی قابلیت‌های پیشینه کتابشناختی را داراست. در خصوص افزودن پیشینه‌های مستند موضوعی به پیشینه‌های مجموعه ذکر این نکته حائز اهمیت است که چون در پیشینه‌های مستند موضوعی به غیر از وجود رابطه میان اصطلاحات هم‌ارز با اصطلاح پذیرفته شده، اصطلاحاتی که با اصطلاح پذیرفته‌شده رابطه سلسله مراتبی یا وابسته دارند نیز ثبت می‌شوند، در هنگام جاسازی آن‌ها در پیشینه‌های مجموعه، می‌بایست عناصر دربرگیرنده اصطلاحات مرتبط با اصطلاح پذیرفته‌شده (غیر از رابطه هم ارز)، از پیشینه حذف گردند. در غیر این صورت، جست‌وجوی اصطلاحات مرتبط، پیشینه‌هایی نامرتبط را بازیابی خواهد نمود، و ریزش کاذب را در نظام جست‌وجوی موتورهای کاوش منجر خواهد شد. به طور کلی، برخی از مهم‌ترین جنبه‌های کاربردی الگوهای پیشنهادی برای تولیدکنندگان پیشینه‌های فراداده‌ای عبارت‌اند از:

الف. امکان پیاده‌سازی پیشینه‌های فرادادهای مبتنی بر دیگر استانداردهای فرادادهای حوزه بافت اطلاعات فرهنگی که زبان نشانه‌گذاری گسترش‌پذیر را بستر نحوی اصلی، یا یکی از بسترهای نحوی خود برگزیده‌اند، بر اساس الگوهای ارائه شده برای طراحی و تولید پیشینه‌های فرادادهای در پژوهش حاضر وجود دارد؛

ب. سازمان‌ها و مراکز ارائه‌کننده خدمات اطلاعاتی در سراسر دنیا می‌توانند پیشینه‌های فرادادهای خود را صرفاً بر اساس الگوهای پیشنهادشده تولید نمایند، و یا این که نسخه‌ای مبتنی بر این روش را تهیه، و در معرض نمایه‌سازی موتورهای کاوش وب قرار دهند؛

ج. با توجه به این که بستر نحوی زبان نشانه‌گذاری گسترش‌پذیر همانند دیگر زبان‌های نشانه‌گذاری و دیگر قالب‌های ذخیره‌سازی، انتقال، و نمایش اشیای محتوایی به هیچ زبان نوشتاری دنیا محدود نیست، بنابراین طراحی و تولید پیشینه‌های فرادادهای با محتوای هر زبانی از جمله زبان فارسی نیز بر اساس روش ارائه‌شده امکان‌پذیر است. در ادامه، الگوهایی برای طراحی راهبردهای جست‌وجو به کاربران نهایی پیشنهاد می‌گردند.

"tagname" "element value" [site:...] [filetype:...]

الگوی سوم. راهبرد جست‌وجوی نمونه برای انجام جست‌وجوهای مبتنی بر عنصر در موتورهای کاوش گوگل و یاهو

بخش نخست الگوی سوم، نام برچسب عنصر هدف را در بر می‌گیرد. نشانه ("") در هنگامی که نام برچسب به صورت عبارتی بود، و یا خاص نمودن جست‌وجو به خصیصه‌ها نیز مد نظر بود، به کار می‌رود. البته برای افزایش دقت جست‌وجو می‌توان نام برچسب را هم پیش از واژه یا عبارت مربوط به ارزش مورد جست‌وجو و هم پس از آن به شرط دانستن شکل دقیق عبارت، و یا استفاده از عملگر (*) به جای واژگانی که کاربر آن‌ها را نمی‌داند اما از وجود آن‌ها در عبارت مطلع است، در نشانه ("") به کاربرد. بخش بعدی، کلیدواژه یا عبارت مربوط به ارزش عنصر را در بر می‌گیرد. بخش سوم زمانی استفاده می‌شود که کاربر علاقه‌مند به محدود کردن نتایج جست‌وجو به وب‌سایت یا دامنه‌ای خاص باشد. و بخش چهارم، برای محدود نمودن نتایج بازیابی به نام پسوند فایل همخوان با نام طرح فرادادهای مبنای تولید پیشینه‌هاست. البته در صورتی که پیشینه‌های مبتنی بر طرح خاصی (مانند مارک ۲۱) مقصود کاربر باشد. نشانه "[]" به معنای اختیاری بودن استفاده از این عملگرهاست.

```
("tagname" AND "element value") AND ("tagname" AND "element value")
[site:...] [filetype:...]
```

الگوی چهارم. راهبرد جست‌وجوی نمونه برای انجام جست‌وجوهای مبتنی بر عنصر به صورت ترکیبی در موتورهای کاوش گوگل و یاهو

با توجه به امکان ترکیب چند عنصر یا ارزش‌های آن‌ها با یکدیگر، الگوی چهارم پیشنهاد شده است. بخشی از این الگو شبیه الگوی سوم است، با این تفاوت که به منظور گروه‌بندی واژه‌ها و عبارات کلیدی به همراه عملگرهای آن‌ها، نشانه "()" درج شده است. عملگرهای "AND/OR/NOT " کار ترکیب را انجام می‌دهند. بهره‌مندی از الگوهای ارائه‌شده در این مقاله، نمایه‌پذیری و پیدانمایی نام‌های برچسب عناصر فراداده‌ای توسط موتورهای کاوش وب، و بازیابی پیشینه‌های مرتبط بر پایه قابلیت جست‌وجوی مبتنی بر عنصر توسط کاربران نهایی را در پی خواهد داشت. بدین ترتیب، بهبود میانکنش‌پذیری میان نظام‌های فراداده‌ای و موتورهای کاوش، به افزایش رضایت کاربران از نظام‌های ذخیره و بازیابی اطلاعات وب منجر خواهد شد.

استفاده از الگوهای اول و دوم برای تولید پیشینه‌های فراداده‌ای توسط طراحان نظام‌های فراداده‌ای به عنوان مبنای اصلی تولید، و یا تهیه نسخه‌هایی دیگر از پیشینه‌ها بر مبنای این الگوها، به سهولت و از لحاظ زمانی مقرون به صرفه است. کاربران نهایی نیز با طراحی راهبردهای جست‌وجو بر اساس الگوهای سوم و چهارم به پیشینه‌های مرتبط‌تری دست می‌یابند. بررسی دیگر روش‌های تولید پیشینه‌های فراداده‌ای با نام‌های برچسب نمایه‌پذیر و با قابلیت پیدانمایی در موتورهای کاوش وب به منظور مقایسه کارایی و سهولت اجرای هر یک، و نیز قرار دادن گزینه‌های مختلف در مقابل تولیدکنندگان پیشینه‌های فراداده‌ای به عنوان پژوهش‌های آینده مطلوب خواهد بود.

کتابنامه

آقاعابدی، زهرا (۱۳۸۹). *بررسی تطبیقی کیفیت نمایه‌سازی و رتبه‌بندی پیشینه‌های فراداده‌ای هسته دوبلین و مارک ۲۱ توسط موتورهای کاوش عمومی*. پایان نامه کارشناسی ارشد. گروه کتابداری و اطلاع‌رسانی. دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران

حریری، نجلا؛ طاهری، سید مهدی؛ فتاحی، سید رحمت‌الله (۱۳۹۲). میانکنش‌پذیری نظام‌های فراداده‌ای و موتورهای کاوش وب: تحولات و رویکردهای جاری. *پژوهشنامه کتابداری و اطلاع رسانی*، ۳ (۲).

شریف، عاطفه (۱۳۸۶). بررسی میزان اثر بخشی عناصر ابرداده‌ای بر رتبه بندی صفحات وب توسط موتورهای کاوش. *فصلنامه کتابداری و اطلاع رسانی*. شماره ۳۸ (تابستان ۱۳۸۶)

طاهری، مهدی (۱۳۸۷). مقایسه کارایی طرح فراداده ای هسته دوبلین و قالب فراداده مارک ۲۱ در سازماندهی منابع اطلاعاتی شبکه جهانی وب. *فصلنامه کتابداری و اطلاع رسانی*، شماره ۴۳ (پاییز ۱۳۸۷)

طاهری، مهدی؛ حریری، نجلا؛ فتاحی، رحمت‌الله (۱۳۸۸). بررسی تطبیقی کیفیت نمایه سازی و رتبه بندی اشیای محتوایی حاوی عناصر فراداده ای هسته دوبلین و مارک ۲۱ توسط موتورهای کاوش عمومی. *فصلنامه کتابداری و اطلاع رسانی آستان قدس رضوی*. شماره ۴۸ (زمستان ۱۳۸۸)

طباطبایی امیری، فائزه السادات؛ طاهری، مهدی؛ فرج پهلو، حسین؛ عصاره، فریده؛ معرف زاده، عبدالعلی (۱۳۹۱). بررسی تطبیقی کیفیت نمایه‌سازی و رتبه‌بندی اشیای محتوایی حاوی عناصر فراداده‌ای دابلین‌کور و مارک ۲۱ مبتنی بر XML در محیط‌های اطلاعاتی پیوسته پویا توسط موتور کاوش گوگل و یاهو. *فصلنامه مدیریت و پردازش*، دوره ۲۷، شماره ۴ (تابستان ۱۳۹۱)

Alexa: The Web Information Company (2011). *Global Top 500*. Retrieved 14 Dec. 2011 from: http://www.alexa.com/site/ds/top_sites?ts_mode=global&lang=none

American Society for Indexing (2012).*Indexing the Web*.Retrieved 7 May 2011 from: http://www.asindexing.org/i4a/pages/index.cfm?pageid=3418

Aqa Abedi, E. (2012), "The effect of syntax on the indexing & ranking of metadata records by the web search engine: a comparative study on MARCXML and DCXML metadata records", unpublished Master's thesis, Science and Research Branch, Islamic Azad University, Tehran (in Persian).

Asadi, S.aied; Jamali, H.amid R. (2004). *Shifts in Search Engine Development: A Review of Past, Present and Future Trends in Research on Search Engines*. Retrieved 7 May 2011 from: http://www.webology.ir/2004/v1n2/a6.html

Bifet, A.lbert; Castillo, C.arlos (2005).*An Analysis of Factors Used in Search Engine Ranking*. Retrieved 7 May 2011 from: http://airweb.cse.lehigh.edu/2005/bifet.pdf

Bray, T.im; Paoli, J.ean; Sperberg-McQueen, C. M.; Maler, E.ve; Yergeau, F.rançois (2008). *Extensible Markup Language* (XML) 1.0 (Fifth Edition): W3C Recommendation 26 November 2008. Retrievd 7 Feb. 2012 from: http://www.w3.org/TR/xml/

Buzzi, M.arina; Andronico, P.atrizia; Leporini, B.arbara (2005).*Accessibility and Usability of Search Engine Interfaces: Preliminary Testing*. Retrieved 7 Feb. 2012 from: http://www.ui4all.gr/workshop2004/files/ui4all_proceedings/adjunct/accessibility/58.pdf

Campex (2012).*Top Search Engines*. Retrieved 7 Feb. 2012 from: http://capmex.biz/resources/top-search-engines

ComScore (2012).*comScore Releases May 2012 U.S. Search Engine Rankings*. Retrieved 7 Feb. 2012 from: http://www.comscore.com/Press_Events/Press_Releases/2012/6/comScore_Releases_May_2012_U.S._Search_Engine_Rankings

Expertrating (2012).*XML Tutorial: Embedding XML in HTML*. Retrieved 7 Feb. 2012 from:http://www.expertrating.com/courseware/XMLCourse/XML-Embedding-HTML-8.asp

Gigee, G.rant (2006).*MARC and MARCXML*. Retrieved 5 Nov. 2011 from:http://threegee.files.wordpress.com/2006/05/marcxml.pdf

Gill, T.oney (2008).*Metadata and the Web: Introduction to Metadata*. Retrieved 5 Nov. 2011 from:http://www.getty.edu/research/publications/electronic_publications/intrometadata/metadata.pdf

Google (2012).*Web Master Tools: Meta tag*. Retrieved 5 Feb. 2012 from:http://www.google.com/support/webmasters/bin/answer.py?answer=79812

Henshaw, R.obin; Valauskas, E.dward J. (2001).*Metadata as a Catalyst: Experiments with Metadata and Search Engines in the Internet Journal, First Monday*. Retrieved 14 Dec. 2011 fromwww.librijournal.org/pdf/1999-3pp125-131.pdf

Kelly (2006).*MARC and MARC XML*. Retrieved 14 Dec. 2011 from http://threegee.files.wordpress.com/2006/05/marcxml.pdf

Kyrnin, J.ennifer (2012). *What Is Pre-Formatted Text?*.Retrieved 14 Dec. 2012from:http://webdesign.about.com/od/htmltags/f/blfaqpre.htm

Lewis, E.dward (2008). *Top Ten Search Engines*.Retrieved 14 Dec. 2011 from http://www.seoconsultants.com/search-engines/

Luk, R.obert; Chan, A.lvin; Dillon, T.haram; Leong, H. V. (2000).*A Survey of Search Engines for XML Documents*. Retrieved 14 Dec. 2011 from: http://www.haifa.il.ibm.com/sigir00-xml/final-papers/Luk/XMLSUR.htm

Luk, R.obert; Leong, H. V.; Dillon, T.haram S.; Chan, A.lvin T. S.; Croft, W. B.ruce; Allan, J.ames (2002). *A Surveyin Indexing and Searching XML Documetns*.Retrieved 14 Dec. 2011 from: http://onlinelibrary.wiley.com/doi/10.1002/asi.10056/full

Mascaro, Michelle J (2004). *The Value of Flexibility in Metadata Schemas*.Retrieved 14 Dec. 2011 from: http://etd.ils.unc.edu:8080/dspace/bitstream/1901/57/1/michellemascaro.pdf

Mohamed, Kh.aledA.f. (2006). *The impact of metadata in web resources discovering*.Online Information Review. 30 (2): 155-167

Qin, J.ian (2000). *Representation and Organization of Information in the Web Space: From MARC to XML*. Retrieved 12 Dec. 2011 from: http://inform.nu/Articles/Vol3/v3n2p83-88.pdf

Quevedo-Torrero, J.esus U.baldo (2004). *IMPROVING WEB RETRIEVAL BY MINING THE HTML TAGS FOR KEYWORDS AND EXPLORING THE HYPERLINK STRUCTURES WEB PAGES*.Ph. D. Dissertation, Department of Computer Science, University of Houston. Retrieved 12 Dec. 2011 from: http://wwwlib.umi.com/dissertations/fullcit/3156028

Safari, M.ehdi (2005).*Search Engine and Resource Discovery on the Web: Is Dublin Core an Impact Factor*. Retrieved 12 Dec. 2011 from: www.webology.ir/2005/v2n2/a13.html

Search Engine Watch (2007). *How Search Engines Work*. Retrieved 12 Dec. 2011 from: http://searchenginewatch.com/article/2065173/How-Search-Engines-Work

Sharif, A. (2007), “Study the effectiveness of metadata elements on web page visibility in

publicsearch engines”, available at: http://eprints.rclis.org/handle/10760/9171#.UHPNcVG94hA (accessed December 7, 2012) (in Persian).

Sokvine, L.loyd (2000). *An Evaluation of the Effectiveness of Current Dublin Core Metadata for Retrieval*. Retrieved 12 Dec. 2011 from: www.vala.org.au/vala2000/2000pdf/Sokvitne.PDF

Tabatabai Amiri F, Taheri S. M., Farajpahlou A, Osareh F, Moarrefzadeh A. “Web Search Engines and Indexing and Ranking the Content Object Including Metadata Elements Available at the Dynamic Information Environments”, Journal of Information Processing and Management. 2012; 27 (4) (in Persian).

Taheri, S. M. (2008). “A comparative study on the efficiency of Dublin Core Metadata Initiative and MARC21 Metadata Format in organizing information available on the World Wide Web”, Library and Information Quarterly. No.: 43 (in Persian).

Taheri, S. M ;.Hariri, N.adjla (2012).*A Comparative Study on the Indexing and Ranking of the Content Objects Including the MARCXML and Dublin Core's Metadata Elements by General Search Engines*. Electronic Library.Vol 30, issue 4

Taheri, S.M. and Hariri, N. (2012), “A comparative study on the indexing and ranking of the content objects including the MARCXML and Dublin Core’s metadata elements by general search engines”, Library and Information Quarterly. No.: 48 (in Persian).

Taheri, S.M., Hariri, N. and Fattahi, S.R. (2014), “Interoperability between metadata systems and web search engines: Current Development and Trends”, Library and Information Research Journal 3 (2) (in Persian).

Turner, T.homas P.; Brackbill, L.ise (1998).*Rising to the Top: Evaluating the Use of the HTML META Tag To Improve Retrieval of World Wide Web Documents through Internet Search Engines* Retrieved 12 Dec. 2011 from: http://cat.inist.fr/?aModele=afficheN&cpsidt=1748620

Woychowsky, E.dmond (2003). XML data islands offer a useful mechanism to display Web form: data. Retrieved 12 Dec. 2011 from: http://www.techrepublic.com/article/xml-data-islands-offer-a-useful-mechanism-to-display-web-form-data/1058668

Yahoo (2012).*What are meta tags?*. Retrieved 12 Feb. 2012 from: http://help.yahoo.com/l/us/yahoo/smallbusiness/promotion/meta/meta-01.html.

Zhang, J.in; Dimitroff, A.lexandra (2004).*Internet search engine's response to metadata Dublin Core implementation*. Retrieved 12 Dec. 2011 from http://portal.acm.org/citation.cfm?id=1142111

Zhang, J.in; Dimitroff, A.lexandra (2005a).*The impact of metadata implementation on Webpage visibility in search engine result (Part II)*. Retrieved 12 Dec. 2011 from: http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6VC8-4BHCBX4-2&_user=10&_rdoc=1&_fmt=&_orig=search&_sort=d&view=c&_acct=C000050221&_version=1&_urlVersion=0&_userid=10&md5=a853d410a866732d3f8ab5dd3217d412

Zhang, J.in; Dimitroff, Al.exandra (2005b).*The impact of Webpage content characteristics on webpage visibility in search engineresult (Part I)*.Retrieved 12 Dec. 2011 from: http://web.simmons.edu/~braun/467/part_1.pdf